DOCUMENT RESUME

ED 478 983 TM 035 128

AUTHOR	Whitehurst, Grover J.
TITLE	The Institute of Education Sciences: New Wine, New Bottles.
INSTITUTION	Institute of Education Sciences (ED), Washington, DC.
PUB DATE	2003-00-00
NOTE	14p.; Paper presented at the Annual Meeting of the American Educational Research Association (Chicago, IL, April 21-25, 2003).
PUB TYPE	Reports - Descriptive (141) -- Speeches/Meeting Papers (150)
EDRS PRICE	EDRS Price MF01/PC01 Plus Postage.
DESCRIPTORS	Agency Role; *Educational Research; Educational Researchers; Elementary Secondary Education; Higher Education; *Institutional Mission; *Research Utilization

ABSTRACT

This paper explains the mission of the Institute of Education Sciences and outlines the activities the Institute is undertaking. It also shares the reflections of the Director of the Institute of Education Sciences on the fit, and sometimes misfit, between the current activities of the education research community and the needs of practitioners and policy makers. The statutory mission of the Institute of Education Sciences is to provide decision makers and the general public with information on: (1) the condition and progress of education in the United States; (2) practices that improve academic achievement and access to education opportunities; and (3) the effectiveness of Federal and other education programs. These three functions are the responsibility of the organizational units of the Institute of Education Sciences: the National Center for Education Statistics, the National Center for Education Research, and the National Center for Education Evaluation and Regional Assistance. Both the statutory mission of the Institute and its conceptual model point it toward applied research. Consequently, the primary focus of the Institute of Education Sciences is on work that has a high consideration of use—practical, applied work that is relevant to practitioners and policy makers. The preponderance of issues identified as high priority areas that the Institute is addressing relate to questions of effectiveness. Such questions are addressed most rigorously with randomized field trials, considered the only sure method for determining the effectiveness of education programs and practices. Although randomized trials are the gold standard for determining what works, they are not appropriate for all questions, and can be enhanced with results from other methods. A look at recent education research has indicated that randomized trials represent only about 6% of the research reported in American Association of Educational Research journals, although they predominate in the "Journal of Educational Psychology." These findings show a mismatch between what education decision makers want from research and what the education research community is providing. The Institute of Education Sciences is working to correct this mismatch through its focus on what works in education. (SLD)

Reproductions supplied by EDRS are the best that can be made
from the original document.

TM TMTR 0179

ED 478 983

American Educational Research Association
2003 Annual Meeting Presidential Invited Session
April 22, 2003

The Institute of Education Sciences: New Wine, New Bottles

Grover J. (Russ) Whitehurst, Ph.D.

Director, Institute of Education Sciences

U.S. Department of Education

This is my second presentation at AERA in my role as the principal research officer in the U.S. Department of Education. When I spoke here last year, I was assistant secretary for the Office of Educational Research and Improvement, otherwise known as OERI. I had been on the job for less than a year, and was busy trying to pour new wine into the old OERI bottle. OERI was on its way out as Congress actively deliberated the reform of Federally funded education research. Those deliberations led to the passage of the Education Sciences Reform Act of 2002. Pay attention to the title of the act. I worked with Congress for a year and a half on the bill. Believe me when I tell you that there was strong bipartisan agreement that education research needed reformation. The bill was signed into law by the President on November 5th of last year. Shortly thereafter, the President appointed me as the first director of the entity created by that legislation, the Institute of Education Sciences. So, I'm back. This time I'm busy trying to create the new Institute bottle and fill it with new wine. Since my appointment is for 6 years this will not be the last time you hear from me, as long as I'm invited to return.

I want to accomplish three things with my remarks today. The first is to have you understand the mission of the Institute. The second is to convey in very broad stokes the activities that are underway that serve the Institute's mission. The third is to share my reflections on the fit, and sometimes misfit, between the education research community's current activities and the needs of practitioners and policy makers as related to the mission of the Institute.

The statutory mission of the Institute is to expand knowledge of education in order to provide decision makers and the general public with information on:

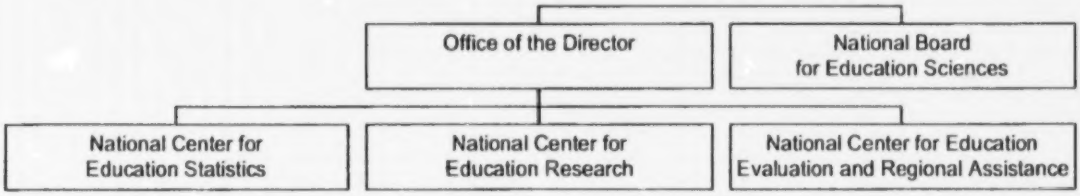
1) the condition and progress of education in the United States
2) practices that improve academic achievement and access to education opportunities
3) the effectiveness of Federal and other education programs

TM035128

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

BEST COPY AVAILABLE

These three functions are the responsibility of the organizational units of the Institute created by statute:


Office of the Director
National Board for Education Sciences
National Center for Education Statistics
National Center for Education Research
National Center for Education Evaluation and Regional Assistance

The National Center for Education Statistics is responsible for gathering and reporting information on the condition and progress of education. The National Center for Education Research is responsible for funding research on practices that improve academic achievement and education opportunities. The National Center for Education Evaluation and Regional Assistance is responsible for evaluating the effectiveness of Federal and other education programs, and for disseminating information to the general public. The Director provides leadership and proposes priorities, which are approved by the National Board for Education Sciences.

Note that under the statute the activities of the Institute are carried out in order to provide useful information to people who have to make decisions about education practices, programs, and policies. In other words, the customers of the Institute are, by law, practitioners, and policy makers, as well as the general public. We are to serve their needs by providing information that will allow them to make better decisions and engage in more effective actions in the realm of education.

Some federal research agencies, by statute, are primarily about the business of basic research and the search for fundamental knowledge. The NSF, for example, has a mission "to promote the progress of science." Other agencies, such as the Institute of Education Sciences, are primarily about practical action, solving real-world problems, and providing useful information to the public at large.

One way of making this distinction is in the terms introduced in the infrequently read but oft cited 1997 book by Stokes, called *Pasteur's Quadrant – Basic Science and Technological Innovation.* Stokes described three categories of research based on two binary dimensions: first, a quest for fundamental understanding, and second, a consideration of use. The work of the theoretical physicist, Niels Bohr, exemplifies the quadrant in which researchers search for fundamental knowledge, with little concern for application. The research of Louis Pasteur,

whose studies of bacteriology were carried out at the behest of the French wine industry, characterizes the work of scientists who, like Bohr, search for fundamental knowledge, but unlike Bohr, select their questions and methods based on potential relevance to real world problems. The work of Thomas Edison, whose practical inventions define the 20th century, exemplifies the work of scientists whose stock and trade is problem solution. They cannibalize whatever basic and craft knowledge is available, and conduct fundamental research when necessary, with choices of action and investment driven by the goal of solving the problem at hand as quickly and efficiently as possible.

Edison's Quadrant (mostly)

		Considerations of Use	
		Low	High
Quest for Fundamental Understanding	Yes	Pure Basic Research (Bohr)	Use-Inspired Basic Research (Pasteur)
	No		Pure Applied Research (Edison)

Each of the scientific quadrants identified by Stokes is important to the common good. Those who argue for the value of basic research have no trouble finding examples of work inspired only by intellectual curiosity that turned out to be extremely practical. Bohrs' work on quantum physics is a case in point.

Without in any way diminishing the value of basic research, whether use-inspired or not, I want to argue for the importance of activities in Edison's quadrant, particularly for topics in which there is a large distance between what the world needs and what realistically can be expected to flow from basic research, and for topics in which problem solutions are richly multivariate and contextual.

Education is such an area: a field in which there is a gulf between the bench and the trench, and in which the trench is complicated by many players, settings, and circumstances. Choose what you consider to be the most exciting developments from basic research in Bohrs' or Pasteur's quadrants that are relevant to education. I'll pick developments in cognitive neuroscience. Paint the rosiest scenario you dare for basic scientific progress in the topic you've chosen over the next 15 years. Then ask yourself what would need to be done to translate those imagined findings into applications that would have wide and powerful effects on education outcomes. I don't know about you, but I'm not optimistic that the results of basic research, even if the findings are powerful, will flow directly and naturally into education. Goodness! Education hasn't even incorporated into instruction what we know from basic research about the effects of massed versus distributed practice – and I learned about that in a psychology course I took in 1962.

Yes, the world needs basic research in disciplines related to education, such as economics, psychology, and management. But education won't be transformed by applications of research until someone engineers systems and approaches and packages that work in the settings in which they will be deployed. For my example of massed versus distributed practice, we need curricula that administrators will select and that teachers will follow that distributes and sequences content appropriately. Likewise, for other existing knowledge or new breakthroughs, we need effective delivery systems. The model that Edison provides of an invention factory that moves from inspiration through lab research to trials of effectiveness to promotion and finally to distribution and product support is particularly applicable to education.

In summary, the Institute's statutory mission, as well as the conceptual model I've just outlined, points the Institute toward applied research, Edison's quadrant.

I've labeled this chart, "Edison's quadrant, mostly," because I understand that it is important to nurture the development of basic knowledge related to education, particularly in areas in which other science agencies and major foundation's aren't involved. Thus, when resources permit, the Institute will support work that examines underlying process and mechanisms, and work that is initiated by the field. For instance, the President's budget request for the Institute for fiscal year 04 includes a healthy amount of money for a field-initiated competition. In addition, many of our new funding programs that are squarely focused on application, such as our program in preschool curriculum evaluation, provide for grantees to carry out parallel research that examines underlying processes.

That said, I reiterate that the primary focus for the Institute will be on work that has high consideration of use, that is practical, that is applied, that is relevant to practitioners and policy makers.

On the issue of relevance, we've recently completed a survey of a purposive sample of our customers to determine what they think we ought to be doing to serve their needs. The sample included school superintendents and principals, chief state school officers, and legislative policy makers. We asked:

> What could the U.S. Department of Education do to make education research more useful, more accessible, or relevant to your work?

Their answers suggest that adjustments are needed in the type of work that is conducted by the education research community. For example, 77% of the school superintendents and local education officials spontaneously criticized existing research for its overly theoretical and academic orientation. A typical response was:

> There may be less than one percent of the existing research that's really meaningful to teachers. Much is for researchers, for getting funding, for career advancement, or for advocacy. . . . I don't want theories. Teachers need strategies, practices. Give them things that can help teaching and learning, things that can help kids.

Another take on the theme of practical relevance emerges from a list of the topics that were identified by respondents as the highest priority issues in need of further research.

1) Effective instructional practices in reading, math, and science
2) Standards and assessment
3) Education finance
4) Closing achievement gaps

Each of these priorities focuses on practical issues about which the customers of education research have to make decisions. They are looking to education research for answers that will enhance the odds that their decisions will be successful. In the context of the requirements of No Child Left Behind and increased public scrutiny of education, they feel they can no longer afford to make decisions based on intuition or opinion. They want to know, for example, how to structure a teacher induction program to enhance retention and teacher performance. They want to know which of the commercially available mathematics curriculum are effective in enhancing student learning. They want to know how to design an assessment and accountability system so that negative effects are minimized. They want to know how they can structure teacher compensation to attract and retain the best and the brightest.

Speaker departs from text to describe evaluation, research, statistics, and dissemination activities of the Institute of Education Sciences – this information can be obtained at http://www.ed.gov/offices/IES/.

The preponderance of the issues that are identified as high priority research areas by our customers and that we are addressing on our evaluation, research, and dissemination programs resolve to questions of effectiveness. In other words, what works best, for whom, under what circumstances? Which preschool programs, or math curricula, or programs for English language learners, or teacher professional development programs, or routes to certification, and so forth are effective?

Questions of efficacy and effectiveness, or what works, are causal, and are addressed most rigorously with randomized field trials. The Institute and I have garnered a fair amount of attention for pushing randomized trials, both in funding programs and in the What Works Clearinghouse. From some quarters the attention has been positive. From others it has been negative. If you have a view on this that is still open, it is important that you understand and form your view based on the Institute's actual position on randomized trials, not a caricature of that position.

This is a synopsis of our position

1. Randomized trials are the only sure method for determining the effectiveness of education programs and practices.

We now have compelling evidence that other methods can lead to estimates of effects that vary significantly from those that would be obtained from randomized trials, nearly always in size and sometimes in direction of effect.

Consider work done by Howard Bloom and colleagues at MDRC (*Can Nonexperimental Comparison Group Methods Match the Findings from a Random Assignment Evaluation of Mandatory Welfare-to-Work Programs?).* The authors compared the findings for a number of non-randomized comparison groups with those for randomized control groups from a large-sample random assignment experiment — the National Evaluation of Welfare-to-Work Strategies (NEWWS). The approach was to generate a non-randomized comparison group for one study, call it study A, from participants who had been in the randomized control group for another study, call it study B. Study B, in turn had a non-randomized comparison group created from study C, and so on. Differences, if any, between results for the randomized control group and results for the non-randomized comparison group for each study were computed. The investigators compared a variety of methods of statistically equating the non-randomized

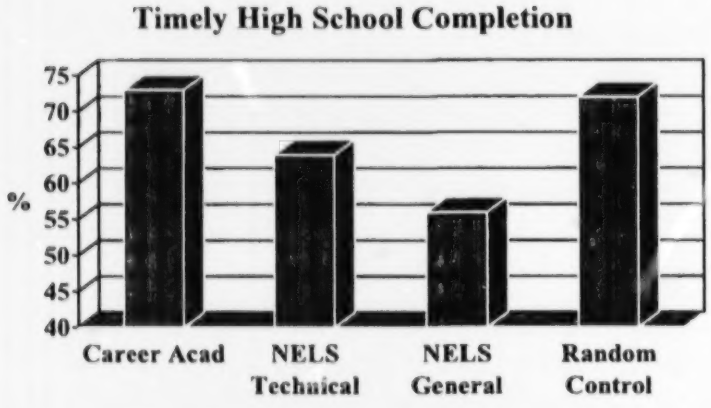
comparison group with the intervention group in each study, such as propensity scores. They also looked at effects for short-term versus mid-term longitudinal outcomes, and for comparison groups formed within the same state as the intervention group versus across state boundaries.

The question, then, is whether quasi-experimental comparison groups, formed with sophisticated statistical methods, generate similar results to control groups formed through randomization. This is what the authors concluded:

> Our results are not encouraging .... For example, three of the five in-state comparison groups produced small biases in the short run while two produced large biases. This suggests that an evaluator using in-state comparison groups to assess a mandatory welfare-to-work program has a 60 percent chance of getting approximately the right answer and a 40 percent chance of being far off. Out-of-state comparison groups performed even less well ....
>
> Adjusting for observed background characteristics did not systematically improve the results. In some cases, these adjustments reduced large biases; in other cases, they made little difference; and in yet other cases, the adjustments made small biases larger. Moreover, there was no apparent pattern to help predict which result would occur.

In other words, quasi-experiments using matched comparison groups have a high chance of producing misleading results, and the most sophisticated statistical matching procedures can increase the chance of error. Those are sobering findings.

Here is a case in point.



The first bar represents high school completion rates for students voluntarily enrolling in a high school technical education program called Career Academies. Data are from participating schools in many locations across the U.S.

The second and third bars illustrate completion rates for students from the National Education Longitudinal Survey who followed a career technical curriculum or a general curriculum in high school. The graph indicates that 73% of the Career Academy group graduated on time versus 64% and 56% of the comparison groups from the NELS. Large study, large N, and pretty impressive results for Career Academies, correct? But the Career Academies study was a randomized trial. The last bar shows the performance of students randomized to the control condition. They graduated at the rate of 72%, not significantly different from the students in the Career Academies intervention.

Randomized trials are the gold standard for determining what works. I've just illustrated why.

2. Randomized trials are not appropriate for all questions.

The development of assessment instruments, for instance, is driven by issues of reliability and predictive validity that are best answered through correlation methods. Questions about the condition and progress of education, the meat of the work of the National Center for Education Statistics, are addressed through surveys, assessments, and data collections, not randomized trials. Efforts to capture in detail the interpretations, beliefs, and circumstances of participants in education are best addressed with narrative and ethnographic methods. Early stages in the

development of new interventions and approaches do not require and can be inappropriately retarded by the use of randomized trials. The use of mathematical modeling to develop and test causal models against large longitudinal databases can be powerful, not as a way to confirm causal hypotheses, but as a way of disconfirming causal models that do not fit the data.

3. Interpretations of the results of randomized trials can be enhanced with results from other methods.

Ethnographies, case studies, surveys, and correlational analyses are all beneficial in making sense of randomized trials that produce variable results across setting and participants, or that produce smaller than desirable effects.

4. A complete portfolio of Federal funding in education will include programs of research that employ a variety of research methods

As I indicated previously, our current and planned research funding at the Institute is consistent with this maxim.

5. Questions of what works are paramount for practitioners; hence randomized trials are of high priority at the Institute.

In summary, randomized trials are one tool in the toolbox. They are to questions of program effectiveness what a hammer is to a nail. You don't use a hammer to saw a board, and you don't use a randomized trial to build a test. But as hammers and nails are essential to carpentry, so are randomized trials and questions of effectiveness at the core of questions that the Institute's customers want research to answer.

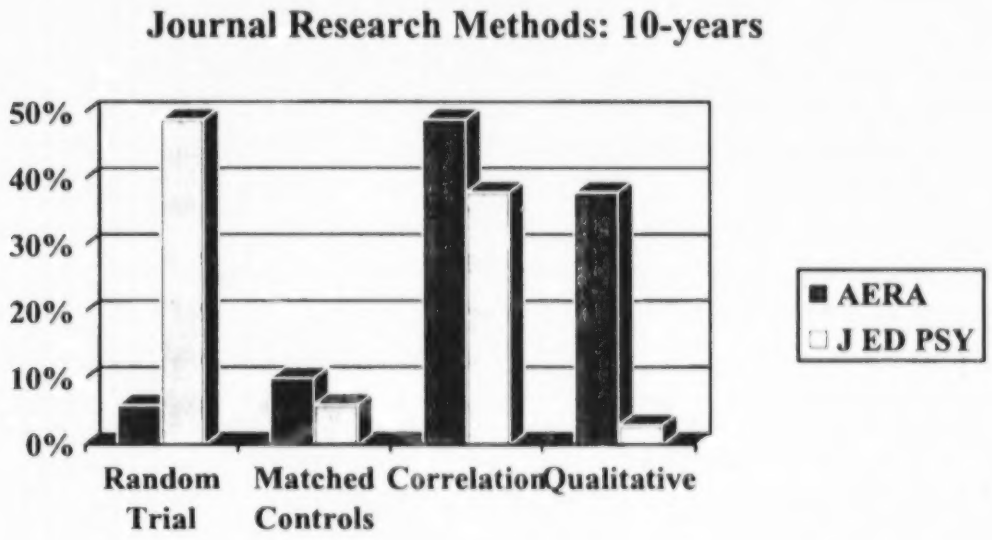
How are AERA and the education research community it represents doing in addressing the research priorities of education practitioners and decision makers, both topically and with respect to randomized trials? The customer survey I described previously suggests that education research is not serving well the practical needs of the field. It is possible, of course, that the administrators and policy makers we surveyed weren't in touch with what is actually going on in education research, or that their knowledge was out-of-date. With the limitations of single sources in mind, I tried to triangulate the current state of the field by considering other sources of data.

I looked through this year's AERA program to identify presentations that seemed to be consistent with high priority, practical questions of the type identified in our customer survey. There are some such presentations, and I applaud them. Other presentations had titles that were

topically relevant but may not have been dispassionate presentations of evidence. Presentations, for example, with titles such as: *No Child Left Behind, Assessment, High Stakes Testing, and Scientifically Based Research: The Axis of Evil.*

Presentations with at least topical relevance to practitioner needs seemed overshadowed by presentations that I expect wouldn't draw the attention of a hard working school superintendent. I'm referring to titles such as *Episodes of Theory-Building as a Transformative & Decolonizing Process: A Microethnographic Inquiry into a Deeper Awareness of Embodied Knowing.*

If you flip through the program, you won't find these exact titles, but you'll find many that are similar.

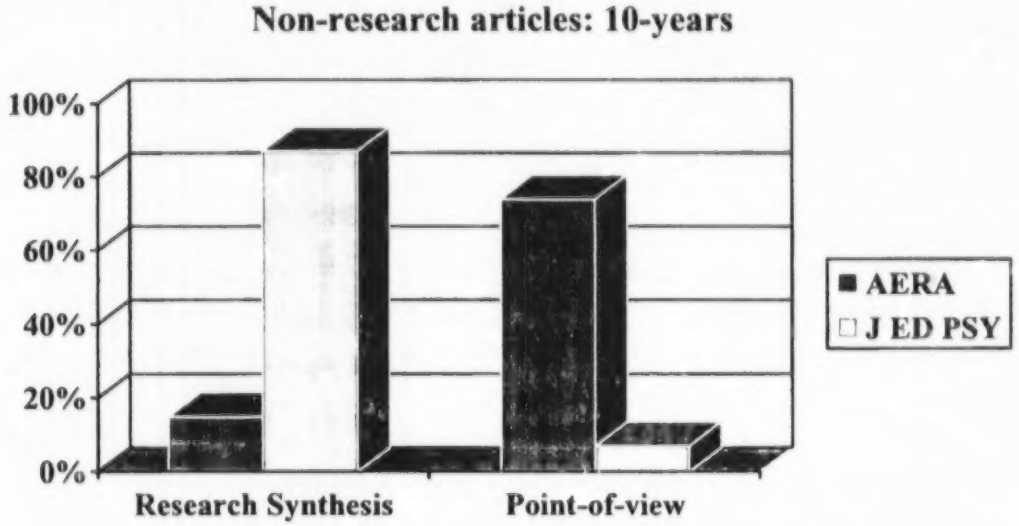


Thinking that a convention program is perhaps not the best source for information on the relative priorities of a scholarly field, I had staff at the Institute examine every article published in AERA's two premier journals, *American Educational Research Journal* and *Educational Evaluation and Policy Analysis*. The examination covered a 10-year span from 1993 to 2002. Articles were first categorized as primary research reports or not. The category of non-research reports included literature reviews, meta-analyses, position pieces, and policy statements. Rejoinders, letters to the editor, and the like were not coded in either category. The research reports were coded into four mutually exclusive categories based on the primary research method

used in the article. The four categories were: randomized trial, matched comparison group, correlational, and qualitative. The chart illustrates the proportion of articles in each category over the 10 years

Only 6% of the research reports in these AERA journals utilized a randomized trial as a primary research method. In contrast, over six times as many studies, used qualitative methods as the primary research tool. If you combine the two categories in which the design is aimed at answering questions of effectiveness -- randomized trials and matched comparison groups -- only 16% of the publications were so designed. Yet what works questions are at the top of the list of research priorities for education decision makers.

Perhaps there is something about education topics that make randomized trials or comparison group designs difficult to apply. To address that possibility, I had articles from the *Journal of Educational Psychology* categorized in the same way and over the same time period as articles from the AERA journals. The results establish that randomized trials predominate in the *Journal of Educational Psychology*. Qualitative studies are as rare there as randomized trials are in the AERA journals.



Even the non-research articles differed substantially between the *Journal of Educational Psychology* and the AERA journals. In the psychology journal, 87% of the non-research articles

were traditional literature reviews or meta-analyses; in both cases the focus was on synthesizing research findings. In contrast, only 19% of the non-research papers in the AERA journals were research syntheses. Instead, 74% of all non-research reports were an expression of a conceptual or political point of view, either an account of the implementation of education policy (usually with suggestions for changes), a review of a concept through a particular theoretical lens, or policy advocacy.

Combining this content analysis of AERA journals, with the content of the AERA convention program, with the feedback we obtained from our survey of customers, I think it would be fair to say that there is a mismatch between what education decision makers want from the education research and what the education research community is providing.

The people on the front lines of education want research to help them make better decisions in those areas in which they have choices to make, such as curriculum, teacher professional development, assessment, technology, and management. These are questions of what works best for whom under what circumstances. These are questions that are best answered by randomized trials of interventions and approaches brought to scale. These are questions and methods and development efforts with which relatively few in the education research community have been engaged.

The people on the front lines of education do not want research minutia, or post-modern musings, or philosophy, or theory, or advocacy, or opinions from education researchers. Recently, a district superintendent asked me what was the best mathematics curriculum for elementary school students. I said there was no research that provided an answer; that all I could offer was my opinion. He said he had enough opinions already. The people on the front lines want to turn to education researchers for a dispassionate reading of methodologically rigorous research that is relevant to the problems they have to solve. They are surrounded by philosophy, and theory, and points of view. They want us, the research community, to provide them a way to cut through the opinion and advocacy with evidence. They feel they aren't getting that.

I have a vision of a day when any educator or policy maker will want to know what the research says before making an important decision. The research will be there. It will be rigorous. It will be relevant. It will be disseminated and accessed through tools that make it useable. The production and dissemination of this research will be in the hands of an education research community that is large, well-trained, and of high prestige. The best and the brightest will understand that there is no more important a task than educating students and no more intellectually challenging and emotionally rewarding a job than to conduct research that meaningfully advances that goal.

I have a vision of a day in which every child receives an education that is good enough, a day in which no child's future is crippled by a bad teacher or a bad curriculum or a bad school, a day in which we have figured out how to deliver an effective education to everyone who wants it. When that day comes, it will be because the nation has learned to ground education practice in science, and when the education research community has learned to engage in a science that serves. I invite you to join the Institute of Education Sciences in that vision and the work that will be required to attain it.

U.S. Department of Education
Office of Educational Research and Improvement (OERI)
National Library of Education (NLE)
Educational Resources Information Center (ERIC)

NOTICE

Reproduction Basis

	This document is covered by a signed "Reproduction Release (Blanket)" form (on file within the ERIC system), encompassing all or classes of documents from its source organization and, therefore, does not require a "Specific Document" Release form.
X	This document is Federally-funded, or carries its own permission to reproduce, or is otherwise in the public domain and, therefore, may be reproduced by ERIC without a signed Reproduction Release form (either "Specific Document" or "Blanket").